“Diesel-related hydrocarbons can dominate gas phase reactive carbon in megacities” by R. E. Dunmore et al.

Response to reviewers

We would like to thank both reviewers for their helpful comments and have addressed specific concerns below.

Anonymous Referee #1

All minor typographical corrections have been completed

OH reactivity should be clearly defined (i.e. lifetime of OH w/r to HC#) and discussion of its importance given a sentence or two.

We have added some additional text to Section 3.2.

“OH reactivity is defined as the total pseudo first order rate coefficient for loss of OH when reacting with VOCs in the atmosphere. This is important in urban atmospheres that are VOC limited, such as London, as the reaction of VOCs with the OH radical is the driving force for the formation of O_3 and other secondary pollutants.”

9543/13 Add references such as: Robinson et al., “Rethinking organic aerosols: Semivolatile emissions and photochemical aging”, Science, 2007.

The Robinson paper has been added.

9544/10 Add references such as: Jimenez et al., “Evolution of Organic Aerosols in the Atmosphere”, Science, 2009. Donahue et al., “A two-dimensional volatility basis set – Part 2: Diagnostics of organic-aerosol evolution”, ACP, 2012.

We have added both the suggested references and thank the reviewer for pointing out our omission.

9544/18 “combined approach” : this statement is a bit vague, please clarify further

The text has been changed to make this section clearer.

“The organic mixture in air is complicated further by the presence of secondary oxygenated products. This requires a combined approach to investigate VOC composition, such as using two different gas chromatography systems.”

9544/20 Please add explicit references to the publications on the field studies to which you refer such as: Gentner et. al., "Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions", Proceedings of the National Academy of Sciences, 2012.

Several relevant references have been added, such as Gentner et al (2012) and Bahreini et al (2012).

545/4 “as well as warm.” Do you mean warm starts? This phrase may not be necessary or should be clarified.

This has been changed to read:

"central London atmosphere experiencing both cold and warm start vehicle emissions."

9554/20 Add references such as: Ensberg et al. "Emission Factor Ratios, SOA Mass Yields, and the Impact of Vehicular Emissions on SOA Formation", ACP, 2014. Bahreini et al., "Gasoline emissions dominate over diesel in formation of secondary organic aerosol mass", GRL, 2012.

These references have been added.

9555/10-12 This is likely to be a good assumption, but please add references to support low SOA from aqueous/GLY sources: e.g. Knote et al., "Simulation of semi-explicit mechanisms of SOA formation from glyoxal in aerosol in a 3-D model", ACP, 2014.

We have made modifications to include additional support

9555/25 Please clarify differences between "real-world conditions" of driving and dynamometer tests

This has been changed to read:

"However, Carslaw and Rhys-Tyler (2013) have recently shown that when vehicles are driven under real-world urban conditions (i.e. different engines loads cause variable catalyst temperatures which can lead to limited effectiveness, as opposed to dynamometer tests where the catalyst is held at optimum), the emissions of NO_x from diesel engines have not been reduced as expected given the new technologies implemented."

Anonymous Referee #2

This reviewer felt that the manuscript was lacking in details. Much of this detail was provided in the supplementary information, with specific SI sections clearly referenced to at appropriate points within the main text. However, we have moved some detailed information from the SI to the main paper to reduce any confusion. We have answered specific questions below.

Major Comments

1. Assumption about the source and composition of emissions: It seems like the authors have assumed a priori, presumably based on the work of Gentner et al. (2012), that tailpipe emissions from diesel vehicles (and even gasoline vehicles) can only be composed of hydrocarbons present in the fuel. Accordingly, they have assumed that only these can contribute to OH reactivity, ozone and SOA production in the atmosphere. Work from the group at Carnegie Mellon University (CMU) (Gordon et al., 2014a,b; Jathar et al., 2014) has showed that unspeciated emissions (that include intermediate volatility organic compounds (Robinson et al., 2007) and possibly are oxygenated and/or incomplete products of combustion (Ensberg et al., 2013)) that are hard to measure using conventional techniques and are unaccounted for in emissions inventories could be very important for SOA production. It is likely that they are also important for ozone production and influence atmospheric OH reactivity. I would encourage the authors to think about the implications of the

CMU work on the results/conclusions in this manuscript. Particularly, I wonder if the CMU work alters the source-resolved results (gasoline versus diesel) of this work.

One of the main advantages of our GCXGC approach is that we ***do*** speciate some of what CMU would class as '*unspeciated*' using our group type analysis. We believe that the use of our lumping technique, achievable due to the second dimension separation (moved from the SI into the main paper as Section 3.1), provides a more accurate and comprehensive analysis of VOC loading. We measure up to an equivalent vapour pressure as a C_{13} hydrocarbon. We find this region to be dominated by alkanes and a small amount of mono and di-aromatics. In the lumping technique used by CMU (Zhao et al., 2014), the "*primary IVOC*" is measured using *m/z* 57, giving a signal essentially representing alkyl fragment ions. This signal is calibrated using the closest *n*-alkane, similar to our approach but in reality this signal is likely from more than just alkanes and the relative proportion of m/z 57 to the TIC was inconsistent between dynamometer and tunnel measurements. In our approach we use a flame ionisation detector, coupled with improved speciation to quantify the diesel alkane fraction with greater confidence.

There is little evidence from the GCXGC of a large amount of OVOCs above 6 carbons, however we have not characterised the losses due to sampling for these species and so they may be being lost during our analysis (Note the CMU study uses offline Tenax tubes and so similar losses are expected – "First, oxygenated compounds only partially elute from the GC column in their underivatized form" Zhao et al., 2014). For this reason, we do not feel the previous data impacts the work presented in our paper. However, we have moved the discussion of our grouping technique and its improved speciation into the main body of the paper, for clarification purposes.

Residual analysis; not absolute analysis: The VOCs measured in this work (especially in summer) are somewhat processed and hence the ozone reactivity and SOA potential assessed in Figures 3, 4, 6 and 7 represent the residual OH reactivity and the residual potential of the sources (e.g., gasoline, diesel, biogenics, etc) to form ozone and SOA. However, the authors seem to suggest otherwise. I would be willing to agree that the analysis holds for winter where very few of the VOCs will have reacted.

We agree with the reviewer that it is likely the VOCs observed in the summer could be classed as the residual reactivity. In winter there is much less photochemistry and so our measurements will be much closer to the source profiles (i.e. the residual is minimal). We have included the term "*residual*" in the discussion of the summer data.

2. No atmospheric loss assumption: I am not sure I understand what the authors mean by the "no atmospheric loss" assumption to deduce the concentration of organics in the C14 to C22 range. The Gentner work develops distributions of organics that are primary in nature, i.e. unoxidized. In contrast, this work measures somewhat-processed organics and hence the distribution of organics will be very different mostly because the higher carbon number species will react faster than the lower carbon number species. Hence, if I have interpreted the authors correctly, I think the assumption about "no atmospheric loss" on Page 9550, line 22 is quite poor. With the current method, not only is the residual reactivity (of what is left in the atmosphere) estimated to be higher but also that sources with higher carbon number compound emissions (e.g., diesel) will appear relatively more important than sources with lower carbon number compound emissions (e.g.,

gasoline). I would encourage the authors to think about ways in which this artefact could be corrected. For instance, the data could be corrected based on the relative [anticipated] differences in reaction rate constants (may be as a function of carbon number). Let's say one emits equal parts of a C10 and C20. If a C20 reacts four times faster than a C10, then after a certain time, one should see four times less C20 than C10.

The reviewer is correct, the "no atmospheric loss" assumption is only really robustly applicable to the winter data. The measurements were taken close to the traffic source and the site is representative of an urban background site. The reduced photochemistry is likely to have lead to minimal losses between emission and observation of the hydrocarbons and so the air mass is likely to be dominated by "unoxidised" components (as is indeed observed). The lack of rate constant data for the higher hydrocarbons makes any meaningful correction impossible at present but in future could allow this to be taken into account. That being said, we have attempted to estimate what impact the higher carbon number species would have. We used the *n*-dodecane (C_{12}) rate constant with OH for the calculated 'unobserved' diesel fraction. By comparing this to the rate constants and lifetimes for C_{16} and C_{20} *n*-alkanes (in the following table), there is only a maximum of a factor of two difference (not a factor of 4 as suggested by the reviewer). From Figure 2 and Table S5 of Gentner et al (2012) it is possible to determine that the majority of the mass emission from the diesel engine tested is below C_{16} and after C_{20} there is a sharp decline; C_7-C_{11} contains 25%, C_{12}-C_{15} contains 34%, C_{16}-C_{19} contains 25% and C_{20}-C_{25} contains 16%, therefore the relative loss of these more minor species will introduce only small overestimates.

Carbon #	OH rate constant (10^{12}) / cm^3 molecule $^{-1}$ s^{-1}	Lifetime with OH[c]
12	13.2[a]	21 hours
16	23.2[a]	12 hours
20	25.2[b]	11 hours
[a]Atkinson and Arey, 2003, Chemical Review		
[b]Calculated using a Structure Activity Relationship as detailed in the Supplementary Information Section 1.8		
[c]Calculated as a 12 hour daytime average with an OH concentration of $1x10^6$		

We have included several caveats throughout the discussion of results, particularly those in summer. These inform the reader that although we are presenting the summer data, this must be used with care as it is likely what we observe is the 'residual' left over after transport and photochemical reactions. As such it is possible that not only do we slightly overestimate the impact of the diesel related hydrocarbons but also underestimate the importance of some species, particularly the OVOCs in summer. However, we believe our analysis to be robust and with the inclusion of the caveats, our results are not misleading.

Minor Comments

Page 9542, line 8: I do not find that the measurement methods provide as comprehensive a speciation as Gentner et al. (2013) or some recent techniques developed by the groups of Allen

Robinson (Zhao et al., 2014) and Robert Yokelson (Yokelson et al., 2013), to name a few. Hence, the claim of a comprehensive speciation does not seem warranted.

The Gentner paper has a very comprehensive characterisation of the liquid fuels and the primary emissions. In the supplementary there is discussion of 200 species measured in the ambient atmosphere but little information is presented. The Yokelson paper describes extensive laboratory emissions characterisation (over 200 VOCS) but only 21 trace gas species are characterised in the field. There are many comprehensive investigations of the urban atmosphere, usually from a very limited number of samples. However our results represent long-term measurements in the urban atmosphere. While we do not speciate an extremely large selection of individual VOCs, the novelty of our approach is in the lumping technique, which allows us to characterise a larger proportion of the total VOC loading than previous studies. Our lumping technique also allows for a better estimate to be made of the functionality of the species included in each group. We believe this to be one of the most (if not the most) comprehensive observations of a UK and specifically London's urban environment. Again this is described in full in the supplementary material. We have moved this section to the main text for clarity and is shown below:

Grouping of unresolved complex mixture

In previous studies using GC-FID, the larger hydrocarbon fraction, where diesel VOC emissions are predominately found, is part of an unresolved complex mixture (UCM). One method used to estimate the relative amounts of VOCs in this region, is to identify the *n*-alkane (which is often observed as a well-defined peak above a raised baseline) and then integrate the area above the blank baseline between two consecutive linear alkanes (using an FID) or to use the *m/z* 57 fragment ion to represent primary IVOC (Zhao et al., 2014). In reality, this gives an estimate of the total or alkyl containing IVOC loading within this volatility range and will not only include the hydrocarbon fraction with that specific carbon number but other compounds as well (*i.e.* lower carbon number aromatics, OVOCs). This study details the improved resolution of VOCs using GC×GC to allow a more stringent grouping of the UCM by carbon number and functionality, rather than by volatility.

Higher carbon number aliphatic compounds (C_6-C_{13}, predominantly alkanes with some alkenes and cycloalkanes), C_4 substituted monoaromatics and C_{10} monoterpenes have been grouped together and the combined class abundance estimated using a response ratio to the corresponding straight-chained *n*-alkane, 1,3-diethyl benzene and α-pinene respectively. The group boundaries are shown in Figure 2, where for example, box 7 corresponds to the C_{12} aliphatic group and encompasses alkanes, cyclic alkanes and alkenes. Only the material within the box is integrated within this retention window. This is a clear improvement over the 1D case, as there are a considerable number of peaks in Figure 2 that would co-elute with the aliphatic group if the entire retention window was co-sampled (*i.e.* aromatics, oxygenates and other hetero species).

Unfortunately, the separation of the linear alkanes, branched alkanes, cyclic aliphatic and alkenes on the GC×GC chromatogram is not sufficient at higher carbon numbers to allow them to be more fully resolved. This is a direct consequence of the use of the cryogen free, and field deployable valve modulator, which when used in total transfer mode, where the flow in the first column slows during the modulation pulse, imposes restrictions on the column dimensions and internal diameters that can be used (Lidster et al., 2011). Also, given the temperature constraints on this instrument, it is likely that the GC×GC not only miss a fraction of the C_{13} aliphatic group but may also be under-reporting the number of isomers in the higher carbon number group. This would explain why the number of isomers decreases after C_{11} aliphatics, rather than increase as would be expected. The aliphatic groups have diurnal behavior (discussed in the next section) that indicates a dominant traffic related source. Fuel composition measurements suggest there is unlikely to be significant quantities of alkenes from traffic related sources; gasoline contains around 3-4 wtC % of alkenes, and diesel contains negligible quantities (Gentner et al., 2012).

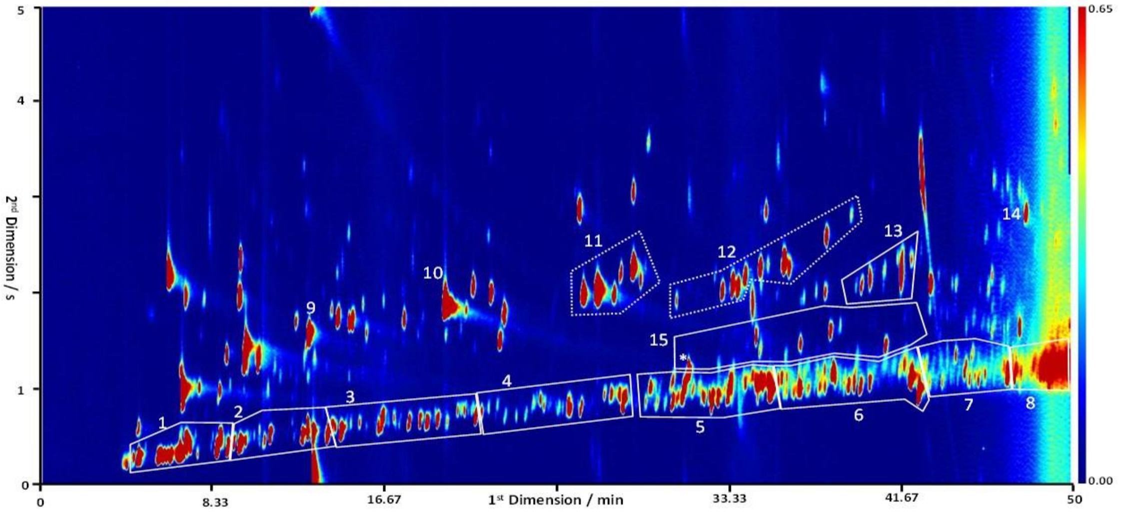

Figure 2: Typical GC×GC-FID chromatogram from 2012-07-25, demonstrating the grouping of compounds. Labelled peaks and groups are identified as follows, with the dashed and solid lines indicating compounds that were identified individually and as a group respectively; (1-8) aliphatic groups from C_6 to C_{13}, (9) benzene, (10) toluene, (11) C_2 substituted monoaromatics, (12) C_3 substituted monoaromatics, (13) C_4 substituted monoaromatics, (14) naphthalene, and (15) C_{10} monoterpenes with * corresponding to α-pinene which is the start of that group. The remaining compounds, not enclosed in a box contain hetero-atoms, primarily oxygenates. The grouping of compounds was accomplished using the lasso technique in Zoex GC Image software (Zoex, USA). This technique allows the software to calculate the area of all the peaks included in the lasso.

Page 9544, line 19 and onwards: The authors talk about the relevance of gasoline and diesel on urban SOA production but fail to cite the relevant references. I would recommend a short summary of what has been found so far. For example, the work of Bahreini et al. (GRL, 2012), and Gentner et al. (2012).

We have included further reference to these papers in the text.

Figure 2: Clearly, the font size on this figure is too small and needs to be increased for better viewability. Same comment applies to Figures 1, 5, 6 and 7 in supplementary material.

The font size has been increased.

Figure 4: The bars can be thinned significantly to create a single column figure.

This has been done

Page 9549-9550, Section 3.2: Could the authors comment about the uncertainty in their estimates (visualized in Figure 3) of OH reactivity based on the assumption of using a surrogate n-alkane to model OH reactivity? Further, it is unclear how the authors separate the estimates in Figure 3 into each source? They cite the work of Gentner et al. (2012) but in doing so are they assuming that emissions of gasoline and diesel are nothing but unburned hydrocarbons?

This information is given in Section 1.8 of the supplementary information where we discuss in detail how we calculated the OH reactivity and explained the use of an *n*-alkane surrogate. We could move this information into the main paper if the editor deems this to necessary, however we believe it would detract from the main focus of the paper. We are not assuming that emissions from fuel are only unburned hydrocarbons rather that this is the only fraction we can measure. For instance, we have not included ethanol as a gasoline source compound as it is an OVOC whose source strengths are less defined due to secondary production. The discussion of source apportionment is detailed in Section 1.1 of the Supplementary Information and Section 3.2 of the main paper.

Section 3.4: Page It is unclear what methods the authors have used to determine emissions of organics by carbon number based on the measurements. I would recommend that the authors describe the methodology in a “Methods” section.

This information can be found in the Supplementary Information. Here we discuss in great detail our methodology. We believe to have this in the main body of the text would distract from the overall message.

Sections 3.5 and 3.6: I would like to see a description of the methods used to the determine ozone potential in Figure 6 (did the authors use the box model of Bill Carter?) and the SOA potential in Figure 7 in a “Methods” section before discussion of the results in Sections 3.5 and 3.6.

We describe the calculation of OFP on Page 9553, lines 6-28. This relatively simple calculation is based on the latest version of the Maximum Incremental Reactivity (MIR) scale determined by Bill Carter using SAPRC-07 model simulations (Carter, W.: SAPRC-07 Atmospheric Chemistry Mechanisms and VOC Reactivity Scales, available at: http://www.engr.ucr.edu/~carter/SAPRC/, 2010. 9553 and Carter, W.P.L., Updated maximum incremental reactivity scale and hydrocarbon bin reactivities for regulatory applications, California Air Resources Board Report, Contract 07-339, 2010.), and so doesn’t require an individual methods section. Again, for the SOA potential, on Page 9555, lines 4-8 we discuss the simple multiplication employed. If the editor deems it necessary we could expand these sections.